An empirical comparison of joint and stratified frameworks for studying G x E interactions

Sung, Yun Ju; Winkler, Thomas W.; Manning, Alisa K.; Aschard, Hugues; Gudnason, Vilmundur; Harris, Tamara B.

Published in:
Genetic Epidemiology

DOI:
10.1002/gepi.21978

Publication date:
2016

Document Version
Peer reviewed version

Table 1: The 20 Participating Cohorts of European Ancestry and their Sample Sizes. Cohorts are divided into two groups (population-based and family-based) and ordered with respect to sample size within each group.

	Cohort	Sample Size	CurSmk			EverSmk		
			Yes	No	% Yes	Yes	No	% Yes
Population	CROATIA-Korcula	456	112	344	24.6%	237	219	52.0%
	CROATIA-Vis	483	141	342	29.2%	277	206	57.3%
	BioMe	1,480	134	1,346	9.1%	441	1,039	29.8%
	CARDIA	1,649	406	1,243	24.6%	693	956	42.0%
	HealthABC	1,662	106	1,556	6.4%	951	711	57.2%
	RS2	1,998	408	1,590	20.4%	1,410	588	70.6%
	AGES	2,410	345	2,065	14.3%	1,440	970	59.8%
	MESA	2,591	298	2,293	11.5%	1,447	1,144	55.8%
	RS3	2,966	673	2,293	22.7%	2,031	935	68.5%
	CHS	2,975	357	2,618	12.0%	1,595	1,380	53.6%
	RS1	4,991	1,162	3,829	23.3%	3,317	1,674	66.5%
	GS:SFHS*	6,439	994	5,445	15.4%	3,133	3,306	48.7%
	ARIC	9,465	2,339	7,126	24.7%	5,685	3,780	60.1%
	WGHS	22,983	2,680	20,303	11.7%	11,284	11,699	49.1%
Family	HERITAGE	499	75	424	15.0%	191	308	38.3%
	GENOA	1,064	169	895	15.9%	535	529	50.3%
	HyperGEN	1,251	114	1,137	9.1%	424	827	33.9%
	ERF	2,491	984	1,507	39.5%	1,721	770	69.1%
	FamHS	3,683	523	3,160	14.2%	1,668	2,015	45.3%
	FHS	8,195	2,520	5,675	30.8%	4,281	3,914	52.2%
	Total	**79,731**	**14,540**	**65,182**	**18.2%**	**42,761**	**36,970**	**53.6%**

Abbreviations: BioMe, Biobank of Institute for Personalized Medicine at Mount Sinai; CARDIA, Coronary Artery Risk Development in Young Adults; HealthABC, Health, Aging, and Body Composition study;RS2, Rotterdam Study cohort 2; AGES, Age Gene Environment Susceptibility Study; MESA, Multi-Ethnic Study of Atherosclerosis;RS3, Rotterdam Study cohort 3;CHS, Cardiovascular Health Study;RS1, Rotterdam Study cohort 1;GS:SFHS, Generation Scotland Scottish Family Health Study; ARIC, Atherosclerosis Risk in Communities;WGHS, Women's Genome Health Study; HERITAGE, Health, Risk Factors, Exercise Training and Genetics;GENOA, Genetic Epidemiology Network of Arteriopathy; HyperGEN, Hypertension Genetic Epidemiology Network; ERF,Erasmus Rucphen Family study; FamHS, Family Heart Study; FHS, Framingham Heart Study

*For this manuscript, GS:SFHS, although a family-based study, removed related individuals using IBS values calculated from genetic data.

Table 2: Correlation between the two frameworks for cohort-specific GWAS results. Scatterplots are shown in Figures 1, 2, and S1.

	Cohort	CurSmk		EverSmk	
		1DF	2DF	1DF	2DF
Population	CROATIA-Korcula	0.943	0.942	0.973	0.950
	CRO-Vis	0.951	0.927	0.970	0.923
	BioMe	0.984	0.990	0.994	0.995
	CARDIA	0.968	0.976	0.996	0.997
	HealthABC	0.993	0.994	0.998	0.998
	RS2	0.992	0.994	0.999	0.999
	AGES	0.997	0.998	0.999	0.999
	MESA	0.977	0.986	0.990	0.993
	RS3	0.998	0.999	1.000	1.000
	CHS	0.991	0.994	0.999	0.999
	RS1	0.996	0.998	0.996	0.997
	GS:SFHS	0.978	0.980	0.995	0.991
	ARIC	0.992	0.994	0.992	0.993
	WGHS	0.999	1.000	1.000	1.000
Family	HERITAGE	0.762	0.819	0.886	0.902
	GENOA	0.998	0.998	0.992	0.992
	HyperGEN	0.885	0.921	0.935	0.942
	ERF	0.973	0.979	0.974	0.979
	FamHS	0.926	0.950	0.960	0.968
	FHS	0.935	0.951	0.939	0.951

Table 3: Spearman rank correlation coefficients between the two stratum-specific genetic effects calculated from the genome-wide resultsused for the 1 DF test in the stratified framework

	Cohort	CurSmk	EverSmk
Cohort-level for population-based cohorts	CROATIA-Korcula	0.000	-0.003
	CROATIA-Vis	0.014	0.005
	BioMe	0.001	0.002
	CARDIA	0.000	-0.002
	HealthABC	0.007	0.010
	RS2	0.003	-0.001
	AGES	0.016	0.014
	MESA	0.012	0.044
	RS3	0.006	0.006
	CHS	0.001	0.004
	RS1	0.013	0.005
	GS:SFHS	0.003	0.006
	ARIC	0.012	0.012
	WGHS	0.014	0.027
Cohort-level for family-based cohorts	HERITAGE	0.105	0.076
	GENOA	0.017	0.030
	HyperGEN	0.052	0.093
	ERF	0.053	0.066
	FamHS	0.071	0.078
	FHS	0.091	0.112
Meta-level	Population-based cohorts	0.034	0.045
	Family-based cohorts	0.090	0.095
	All cohorts	0.055	0.065

Table 4: Correlation between the two frameworks for meta-analysis results. Scatterplots are shown in Figures 4 and 5.

Stratified Framework with	Meta-analysis with	CurSmk		EverSmk	
		1DF	2DF	1DF	2DF
Scheme A	Population cohorts	0.942	0.970	0.950	0.982
	Family cohorts	0.860	0.893	0.889	0.924
	All cohorts	0.904	0.947	0.927	0.965
Scheme B	Population cohorts	0.957	0.990	0.965	0.995
	Family cohorts	0.882	0.946	0.905	0.950
	All cohorts	0.923	0.98	0.948	0.985

Table 5: Comparison of schemes A and B for family-based meta-analysis forCurSmk at select variants

Marker	Level	Type	N E=0	MAC E=0	MAC E=1	Effect E=0	StdErr E=0	P E=0	Stratified 2DFP	Interaction 2DFP
2:48619812 (MAF=1.2%)	Meta	Scheme A	4,479			12.8	1.0	7.8E-41	8.5E-40	
	Meta	Scheme B	3,160			0.6	2.7	0.83	0.63	0.89
	Cohort	FamHS	3,160	77.7	13.7	0.6	2.6	0.83	0.60	0.89
	Cohort	GENOA	895	29.9	<10	3.2	3.1	0.30		
	Cohort	HERITAGE	424	10.3	<10	15.8	1.0	1.2E-54		
6:142093034 (MAF=1.7%)	Meta	Scheme A	12,798			4.7	0.5	3.8E-23	2.2E-22	
	Meta	Scheme B	10,342			-0.1	1.0	0.89	0.47	0.61
	Cohort	ERF	1,507	66.9	52.7	1.2	2.1	0.58	0.68	0.78
	Cohort	FamHS	3,160	140.3	25.2	0.2	1.6	0.88	0.98	0.89
	Cohort	FHS	5,675	141.8	63.0	-0.5	1.5	0.74	0.47	0.5
	Cohort	GENOA	895	43.8	<10	-4.0	2.6	0.12	0.15	0.12
	Cohort	HERITAGE	424	19.1	<10	-6.3	0.5	1.9E-33		
	Cohort	HyperGEN	1137	30.5	<10	0.3	4.4	0.95		
12:5679139 (MAF=1.3%)	Meta	Scheme A	4,479			1.9	2.1	0.37	1.8E-09	
	Meta	Scheme B	3,160			4.1	2.8	0.15	5.0E-09	1.9E-11
	Cohort	FamHS	3,160	84.1	12.5	-4.1	2.8	0.14	6.3E-10	3.8E-12
	Cohort	GENOA	895	33.5	<10	3.1	3.7	0.41	0.2	0.21
	Cohort	HERITAGE	424	10.5	<10	-3.6	4.8	0.46	0.59	2.4E-08